

Open Rent Initiative Data Methodology

Last revised: July 8, 2026

Contents

1. Introduction	2
1.2 Document Purpose and Scope	2
1.3 Foundational Approach	2
2. Technical Infrastructure	4
2.1 Data Pipeline Overview	4
2.2 Tooling and Stack	4
2.3 Open Source and Open Data Commitments	5
3. Foundational Property Data	6
3.1 Data Sources for Foundational Property Inventory	6
3.2 Property Classification Framework	7
3.3 Unit Classification Framework	8
3.4 Address Standardization and Entity Resolution	9
3.5 Unit Identity and Deduplication	10
4. Rental Data	11
4.1 Rent Basis: Asking vs. In-Place	11
4.2 Rent Source Taxonomy	11
4.3 Data Sources	12
4.4 Affordable Rent Calculation (TCAC Method)	13
4.5 Multi-Source Reconciliation	13
5. Coverage and Limitations	15
5.1 Geographic Scope	15
5.2 What Open Rent Observes and Does Not Observe	15
5.3 Known Coverage Gaps	15
5.4 Appropriate and Inappropriate Use	16
6. Additional Methodologies	17
6.2 Current Addenda	17

1. Introduction

The Open Rent Initiative is a civic-technology non-profit project created by Citizen Codex and fiscally sponsored by the PSL Foundation. Open Rent aims to make the fragmented landscape of rental housing more easily accessible. The initiative is supported by grants from the Silicon Valley Community foundation, Chan Zuckerberg Initiative and the San Francisco Foundation, whose investments have enabled the build-out of the platform's data infrastructure and its first municipal pilot.

By aggregating parcel, government housing data, and market data into a common, transparent framework, Open Rent generates the evidence base needed to view the holistic rental composition of a given geography and assess the impact of housing programs and policy.

1.2 Document Purpose and Scope

This document describes the data sourcing, processing, and maintenance methodology underlying the Open Rent platform. It is intended to outline how Open Rent constructs its rental housing inventory and rent figures, what evidence underlies each data point, and where the data has known limitations.

Open Rent's first pilot deployment is in the city of Mountain View, California, a choice grounded in both opportunity and need. California has among the nation's most acute affordability pressures and one of its most layered rent-regulation environments, making it a demanding proving ground for any rental data infrastructure. Santa Clara County combines extreme rent levels driven by the regional tech economy with a diverse mix of housing stock, offering the full range of property types the platform is designed to characterize.

Lessons from Mountain View are now informing expansion into other selected geographies, where the existing framework is translated to different housing markets and regulatory contexts.

The pipeline architecture is designed for multi-jurisdiction expansion. This methodology document will be updated and published with each new release. This document does not cover analytical interpretations of the data, platform interface decisions, or policy conclusions.

1.3 Foundational Approach

Open Rent constructs its rental housing database through three sequential layers, each dependent on the one preceding it.

The first layer establishes the foundational property inventory: what physical structures exist in a given city, how many units each contains, and how each property is characterized

by use, age, and legal status. Without this layer, there is no canonical record against which rental status and pricing can be attached.

The second layer determines the regulatory status of each property and each unit within it: whether the property is enrolled in a government program, subject to a rent stabilization ordinance, or restricted by an affordability covenant, and what class of restriction applies to each individual dwelling. This layer produces the property and unit classification used throughout the platform.

The third layer captures what each unit rents for and on what evidentiary basis: whether the figure comes from an official registry filing, a market listing, a regulatory formula, or a statistical model. The appropriate method for deriving a rent figure depends on the regulatory classification assigned in the second layer; the third layer cannot be applied without it.

2. Technical Infrastructure

2.1 Data Pipeline Overview

Open Rent processes data in four layers. Each layer cleans, standardizes, or enriches the data a little more than the last, and at every step the original source record remains intact, allowing any number on the platform to be traced back to exactly where it came from.

Raw (Bronze). Each external source is ingested with minimal transformation. Every row at this layer carries two audit columns: a pipeline run identifier (`audit_run_id`) and a load timestamp (`audit_loaded_at`). These columns enable full lineage from any production figure back to its source record and the pipeline run that ingested it.

Silver. Source-specific normalization is applied at this layer. Addresses are standardized to a canonical form (described in Section 3.4). Geographic coordinates are assigned or verified according to a source priority hierarchy. Preliminary within-source deduplication removes records that represent the same physical unit under different identifiers within a single source.

Gold. Entity resolution produces canonical property and unit records. Addresses from multiple sources are clustered into canonical property identities. Unit records from multiple sources are resolved to single physical slots. Regulatory classification is applied based on program enrollment and registry status. The gold layer is the authoritative source for all downstream queries.

Views and API. Two materialized database views aggregate the gold layer for consumer queries. `v_unit_detail` exposes one row per unit with both asking and in-place rent values and the unit's regulatory classification. `v_property_summary` exposes one row per property with summarized unit counts and property-level characteristics including building type, year built, and ZIP code.

2.2 Tooling and Stack

Pipeline orchestration is handled by Dagster, which manages asset dependencies, run schedules, and execution history. The primary data store is PostgreSQL with the PostGIS extension for spatial geometry operations. Schema migrations are managed via Alembic with a single-head constraint enforced at every deployment, preventing migration divergence across branches or environments.

Pipeline logic is written in Python. The web platform is built on SvelteKit and TypeScript, using Drizzle ORM for type-safe database queries. Source records carry audit columns used for lineage tracing and incremental processing; skip cursors at the asset level ensure that unchanged source data does not trigger redundant reprocessing.

2.3 Open Source and Open Data Commitments

Open Rent maintains a strong commitment to transparency, collaboration, and the principles of open research around rental data. For those interested in gaining further access to the data or to discuss the architecture and methodology please reach out.

3. Foundational Property Data

Foundational property data serves as the canonical inventory of rental properties in a city: what structures exist, how many units they contain, and their physical and legal characteristics. All subsequent layers, such as regulatory classification and rent pricing, are attached to records produced in this layer.

Three terms carry specific meanings throughout this document. A property is the legal contracting entity; it may span multiple buildings and addresses. A building is a single physical structure at one address. A unit is the smallest discrete dwelling with its own lease. Rent is always set at the unit level, though building and property characteristics influence it.

3.1 Data Sources for Foundational Property Inventory

Four external sources contribute to the foundational property inventory. Each is described below with its content, access method, and known limitations.

Source	Contents	Access Method	Known Limitations
Regrid (CA Santa Clara County)	Parcel geometry (WGS84), LBCS land use codes, structure use description, year built, unit counts, building type, assessed values	Parquet file delivered via S3; loaded in full on each ingestion cycle	Unit counts are assessor-derived and may not reflect actual physical units; geometry may lag recent construction or demolition
Santa Clara County Assessor PDFs	Situs address, mailing address, document type (deed, lien, etc.), transfer date, assessed tax values	Gathered from county assessor website; PDFs cached to S3 to avoid redundant fetches; multi-threaded parser extracts structured fields	PDF availability depends on county record completeness; parse failures logged separately; coverage not guaranteed for all APNs
HUD LIHTC (Low Income Housing Tax Credit)	Unit counts by bedroom size, tax credit type (4% / 9%), income ceiling (% AMI), year placed in service, subsidy program flags (Section 8, HOME, etc.)	National CSV database published by HUD; filtered to California at ingestion; updated on federal LIHTC publication cycles (annually)	National database may lag local project status; unit counts may not match current physical state after renovation or conversion
Mountain View Below -Market- Rate (BMR) list	Unit-level list of below-market-rate units; includes in-place rents, AMI designations (50% and 80%), and a per-unit slot identifier	City-provided CSV; ingested on receipt of updated file from the City of Mountain View	Coverage limited to units enrolled in the city's BMR monitoring program; units with lapsed restrictions or unreported vacancies may be absent

Sources that contribute rent data in addition to property characteristics (specifically HUD LIHTC and the Mountain View BMR list) are also described in Section 4.3.

3.2 Property Classification Framework

Every rental property in Open Rent receives exactly one property classification. This classification describes the regulatory and subsidy status of the property as a whole and determines which rent derivation methods apply to its units. Classifications are mutually exclusive and exhaustive; if a property's program profile satisfies more than one definition, the most restrictive classification is assigned.

Classification	Code	Subsidy?	Means-Tested?	Rent Setting
Locally Unrestricted	market_unrestricted	No	No	property owner sets and adjusts freely. Subject only to state law (AB 1482).
Stabilized	market_stabilized	No	No	Annual increases capped by local ordinance for all units. Rent resets to market at vacancy.
Mixed Income	mixed_income	Some units	Some units	Market-rate units alongside restricted units (inclusionary and/or subsidized).
Affordable	affordable	Yes	Yes	90% or more of units have rents set by publicly subsidized programs, often with ceilings based on AMI.

Classification is derived from the property's program enrollment list using the following decision logic:

- In a subsidized housing database (LIHTC, Section 8)?
 - Yes, with deeply-affordable programs (Sec 8, Public Housing, 811, USDA) applying to all or some units → Affordable
 - Yes, income-restricted only → Affordable
- No, but in BMR / Inclusionary registry → Mixed Income
- No, but in Rent Stabilization Registry → Stabilized
- None of the above → Locally Unrestricted

Two clarifications on this classification are necessary. First, Stabilized is not a subtype of Affordable. Rent stabilization is not means-tested; there is no income qualification for a tenant to rent a stabilized unit and there is no government subsidy paid to the property owner. Where a property appears in both a rent stabilization registry and a subsidized housing database, it is classified as Affordable.

Second, the distinction between Affordable and Mixed Income turns on the share of restricted units. A property where 90% or more of units carry subsidy or deed restrictions is classified as Affordable; a property with a lower share of restricted units among otherwise market-rate units is Mixed Income. Common Mixed Income mechanisms include inclusionary zoning set-asides (typically 10 to 20% of units under local ordinance), density bonus agreements under California Government Code Section 65915, and deed-restricted below-market-rate units within market-rate buildings. During the Mountain View pilot, only the deed-restricted BMR units were included in this classification.

Building type is a separate property-level attribute distinct from classification. It describes the physical structure type and is derived from Regrid land use codes and SCC Assessor records. Values are: Single Family, Condo/Townhouse, Duplex, 3–4 Unit, Small Multifamily (5–49 units), Large Multifamily (50+ units). Building type is used for filtering and disaggregation on the platform and does not affect regulatory classification.

Two edge cases are handled explicitly. Owner-occupied duplexes with one rental unit are classified as Locally Unrestricted with a total unit count of one. Multi-building developments covered by a single program contract are classified as a single property; the classification applies to the contracting entity, not to each building address independently.

3.3 Unit Classification Framework

Every unit within a rental property receives exactly one unit classification. While property classification describes the regulatory status of the contracting entity, unit classification describes the rent restriction in place for that specific dwelling at the point of a lease. For Locally Unrestricted, Stabilized, and Affordable properties, all units share the property classification. Complexity arises only in Mixed Income properties, where market-rate and restricted units coexist within the same building.

Classification	Code	Rent Setting	Vacancy Behavior
Locally Unrestricted	market_unrestricted	No ceiling. property owner sets and adjusts freely. In California, AB 1482 applies a state-level backstop (5% plus CPI, maximum 10%) for properties 15 or more years old.	Resets freely at vacancy.
Stabilized	market_stabilized	Annual increases capped by local ordinance (CSFRA in Mountain View). Initial rent at move-in is market-set; regulation constrains only the growth rate for existing tenants.	Vacancy decontrol applies under the 1995 Costa-Hawkins Rental Housing Act: rent resets to market rate when the unit becomes vacant.
Inclusionary	inclusionary	Rent capped via deed restriction. Set at 30% of gross monthly income at the target AMI tier. Annual increases follow HUD AMI adjustments. No direct government payment to the property owner; market-rate units in the same building cross-subsidize the restricted units.	Restriction follows the unit, not the tenant. Does not reset at vacancy. California minimum restriction period for rental inclusionary units is 55 years.
Subsidized	subsidized	Rent below market through publicly financed programs (LIHTC, Housing Choice Vouchers, HOME, etc.). Set as a percentage of AMI tier or as a share of tenant income. There is a direct financial relationship between a government program and the unit's affordability.	Regulatory agreement or deed restriction persists. Unit re-rents at restricted rate to an income-qualified household. Vacancy decontrol does not apply.

Two distinctions merit emphasis. An inclusionary unit resets the occupant but not the restriction at vacancy; a stabilized unit resets to market rate. A subsidized unit carries a direct government subsidy relationship; an inclusionary unit does not. Here, it's below-market status is produced entirely by a deed restriction, with the property owner bearing the cost through cross-subsidy from market-rate units in the same building.

Where a unit carries both a deed restriction and a government subsidy contract (common in layered affordable housing programs), it is classified as subsidized. The Subsidized classification takes precedence, and the inclusionary designation is reserved for units where the deed restriction is the only active restriction.

Mountain View CSFRA coverage. Not all residential units in Mountain View are subject to rent stabilization. The Community Stabilization and Fair Rent Act (CSFRA), enacted November 8, 2016 under Article XVII of the City Charter, applies selectively based on building age and type.

Note: As Open Rent expands to new cities, each will bring its own stabilization framework and the requirement for dedicated data handling. The platform is designed to accommodate this variation through a rules engine that stores jurisdiction-specific logic as configuration, keeping the underlying data model consistent across cities.

3.4 Address Standardization and Entity Resolution

Records from multiple sources must be resolved to a single canonical property identity before classification and rent data can be attached. This resolution is accomplished through a two-step process: address standardization followed by cluster-based entity resolution.

Address standardization is applied to every address string ingested from any source. A four-tier hierarchical normalization pipeline produces a pipe-delimited five-field canonical string in the form `LINE1|LINE2|CITY|STATE|ZIP5`. Tiers are applied in order, with each successive tier serving as a fallback when the prior tier fails:

1. Pre-normalized. The caller asserts that the address parts are already standardized; no further processing is applied.
2. Strict. The scourgify conditional random field (CRF) tagger is applied with full locality context (city, state, and ZIP code). This tier produces the highest-quality normalization and is the default for addresses with complete locality information.
3. Loose. Scourgify is applied with city and state only, without ZIP code. USPS Publication 28 street suffix abbreviations are applied regardless of tier.
4. Parsed. The usaddress parser is used with USPS Publication 28 abbreviation applied to the output. This tier handles addresses that the CRF tagger cannot parse.

Addresses that cannot be processed by any of the above tiers are stored as raw passthrough strings in uppercase and are flagged explicitly as un-normalized. These records are excluded from clustering until they can be corrected.

Entity resolution groups normalized address records into canonical property clusters. Records sharing the same normalized address string are joined into connected components using a union-find algorithm. Each cluster is assigned a stable identifier equal to the minimum UUID in the component. A canonical address representative is selected from each cluster by geocode source priority (listing data over census coordinates over parcel polygon centroid). Manual override records, maintained as a CSV-loaded table, correct known cases where identical physical locations produced non-matching normalized strings (for example, street name abbreviation variants for the same thoroughfare).

Total unit count for each property follows a precedence hierarchy at cluster creation time: (1) affordable program unit count, (2) rent registry per-address count, (3) Regrid derived unit count. Building type is assigned from Regrid first, then SCC Assessor records, then SFR heuristic.

3.5 Unit Identity and Deduplication

When the same physical unit appears in more than one source database, it must resolve to a single record rather than producing duplicate entries. Open Rent uses a deterministic identifier scheme to accomplish this.

Each physical unit slot is assigned a stable UUID derived from two inputs: the canonical property cluster identifier and a normalized form of the unit's address line 2. Address line 2 normalization strips unit-designator prefixes (such as: apt, apartment, unit, unit, and the # symbol) so that the same physical slot appearing under different source formatting conventions collapses to the same canonical form. For example, "APT 2B," "Unit 2B," and "#2B" all normalize to "2B" and therefore map to the same unit identifier. Designators that describe genuinely distinct space types (ste, fl, rm, lot, ph, bldg, spc) are preserved unchanged, as these typically indicate different units rather than formatting variants of the same unit.

Units without a street address line 2 (that is, units where the source provides only a property address) receive an identifier that incorporates the source name and source record identifier, ensuring that distinct records from different sources remain distinct rather than collapsing into a single slot.

Synthetic placeholder units are created for properties that have a known unit count from assessor data but no individual unit records from any listing or registry source. These placeholders carry a deterministic identifier based on an ordinal position within the property and are replaced by real unit records when listing or registry data arrives. Synthetic units are stable across pipeline runs: an ordinal-1 placeholder at one run remains the same record at the next run unless the unit count decreases, in which case the highest-ordinal placeholders are retired.

When two sources contribute descriptions of the same physical slot, both descriptions are attached to the single unit record. The most restrictive regulatory classification present across all descriptions is applied to the unit.

4. Rental Data

Rental price data is collected through three complementary channels: active listing observation for market-rate asking rents, official registry and program data for in-place and deed-restricted rents, and regulatory formula derivation for income-restricted unit maximums. No single source covers the full spectrum from locally unrestricted market rents to deeply subsidized affordable housing, and so a multi-source approach is required. Every rent figure in Open Rent carries two labels that together fully describe any price in the system: rent basis and rent source tier.

4.1 Rent Basis: Asking vs. In-Place

Rent basis is a required tag on every rent record. Two mutually exclusive values are used.

Basis	Code	Definition	Answers the question
Asking Rent	asking	The rent a property owner would charge a new tenant today: the market-entry price at the time of listing or program enrollment.	What would it cost a new tenant to move into this unit?
In-Place Rent	in_place	The legally binding rent a current tenant is actually paying under their lease: the contracted price in effect as of the most recent registry filing.	What are current tenants in this unit actually paying?

The distinction between these two bases is analytically significant. Asking rents reflect current market conditions at the time a unit becomes available. In-place rents reflect the stock of active leases, which in stabilized buildings can diverge substantially from market asking rents when long-tenured residents have received only the allowable annual adjustments over many years. Both values are stored and exposed independently on the platform; neither is substituted for the other, and neither is derived from the other.

Rent basis applies to every rent record regardless of property or unit classification. A stabilized building, for example, will have both asking rents on record from its initial registry filing and in-place rents from annual update filings.

4.2 Rent Source Taxonomy

Rent source describes where a given price came from and its evidentiary quality. Four mutually exclusive tiers are defined, listed in priority order. When multiple sources provide a rent figure for the same unit, the higher-ranked source is used for the canonical display value. Per-unit labeling accompanies every rent figure on the platform: the source tier and the specific source or method name are both shown.

Tier	Name	Description	Typical Basis	Sources
T1	Actual	From a partner, official filing, or government record. Known data collected by a trusted source. Not inferred, modeled, or derived from a formula.	in_place	MV Rent Registry, MV BMR List, HUD Section 8 Picture of Housing

Tier	Name	Description	Typical Basis	Sources
T2	Listed	Observed rent on a live rental listing platform. Applicable to any unit type or classification. Reflects the advertised price at the time of observation.	asking	Dwellsy, RentCast
T3	Calculated	Computed from known program rules: a formula applied to published regulatory tables. Not a statistical estimate; the inputs and method are fully specified.	asking	CA TCAC rent tables, HUD LIHTC dataset
T4	Estimated	<i>Modeled or statistically inferred. Must explicitly declare whether the output is asking or in-place rent; a model cannot produce both simultaneously.</i>	<i>Declared at model build time</i>	<i>Future Releases</i>

The two dimensions — basis and source tier — are independent. Any combination is valid, though some pairings are more common than others. Actual rents are most often in-place (registry filings reflect what tenants are paying), though an initial registry filing at move-in represents an actual asking rent. Listed rents are almost always asking rents. Calculated rents are typically asking rents representing regulatory ceilings. Estimated rents declare their basis explicitly at model build time.

4.3 Data Sources

Open Rent addresses the fragmented housing data landscape by combining five sources across three categories: public listing platforms, municipal registries, and federal housing program data.

Source	What It Measures	Limitations
Dwellsy / RentCast / RentCafe	Rental listings observed on active listing platforms. Provides asking rent and bedroom count for market-rate and some stabilized units at time of listing.	Coverage limited to actively listed units. Long-term tenanted units are not represented. Listing accuracy depends on property owner-submitted data.
Mountain View CSFRA Rent Registry	Filings by property owners of covered units under the Community Stabilization and Fair Rent Act. Initial filing records rent at the time of tenant move-in; annual updates record the rent currently charged.	Limited to covered units (pre-1995 multifamily, 3+ units). Non-compliant property owners who have not registered are absent. CSFRA exemptions (see Section 3.3) produce gaps in stabilized unit coverage.
Mountain View BMR / Inclusionary List	City-administered registry of below-market-rate units. Provides current in-place rents and AMI designations at the 50% and 80% of area median income tiers.	Coverage limited to units enrolled in the city's BMR monitoring program. Units with lapsed restrictions may not be removed promptly.
HUD LIHTC Dataset	National Low Income Housing Tax Credit database. Provides program-level data enabling calculation of maximum allowable rents via TCAC tables (see Section 4.4).	Derived figures represent regulatory ceilings; actual charged rents may be lower. National database may not reflect current status of individual projects.

Two sources are used for internal coverage validation and benchmarking but do not contribute rent figures displayed to platform users: ACS tables B25003 (tenure by occupancy) and B25024 (units in structure) from the Census Bureau, and the Zillow

Observed Rent Index (ZORI). These sources are used to assess the completeness of Open Rent's unit inventory and to contextualize platform rent figures against independent estimates.

4.4 Affordable Rent Calculation (TCAC Method)

Open Rent derives asking rent figures using the California Tax Credit Allocation Committee (TCAC) rent tables for income-restricted units, which are those whose rent is set by a regulatory formula rather than by market negotiation or registry filing.

LIHTC developers are contractually required under their tax credit regulatory agreements to cap rents at specific AMI levels. TCAC publishes the legal maximum allowable rent by county, AMI percentage, and bedroom size each year. This is not an estimate; it is the regulatory ceiling published by the state and enforceable under the regulatory agreement. If the county, AMI tier, and bedroom count are known for a property, the TCAC maximum can be looked up directly from the published tables.

The derivation proceeds as follows: (1) bedroom mix is obtained from the LIHTC dataset; (2) AMI tier is obtained from the LIHTC income ceiling field; (3) the TCAC rent table for the applicable county and year is queried; (4) a discount factor is applied. The discount factor reflects the observation that many LIHTC properties in Mountain View charge below the TCAC ceiling, typically due to additional local or state subsidies layered on top of the tax credit program. Three discount factor values are in use: 1.00 (no discount), 0.90, and 0.79, calibrated from Mountain View listing data. These factors are specific to California and must not be applied to other jurisdictions without local calibration.

A distinct method applies to deeply-affordable units, where rent is set not by a formula but at 30% of the tenant's actual income, with a government program paying the difference. For these units — primarily Section 8 project-based and public housing — the rent source priority is: HUD Picture of Subsidized Households actuals, then Multifamily Section 8 contract rents, then TCAC table fallback (flagged at lower confidence).

Income-restricted vs. deeply affordable. In income-restricted programs (LIHTC, HOME, State HFA financing), rent is capped at 30% of a specific AMI tier for the unit size. The ceiling is a fixed formula; actual tenant income does not change the rent. In deeply-affordable programs (Section 8, Public Housing, Project-Based Vouchers), rent is set at approximately 30% of the tenant's actual household income and the government pays the property owner the difference. Both are classified as subsidized at the unit level; the difference is captured in the `rent_setting_mechanism` field.

4.5 Multi-Source Reconciliation

When the same physical unit appears in more than one source database, the deterministic unit identity scheme described in Section 3.5 ensures that both source descriptions attach to a single canonical unit record rather than producing duplicate entries.

Both asking and in-place rent values are preserved on the merged record. They are not averaged, collapsed, or replaced; each is displayed independently on the platform with its own source label. When multiple sources contribute a value for the same basis (for example, both a registry filing and a listing for the same unit's asking rent) the higher-ranked source tier determines the canonical value.

Regulatory classification follows the same precedence logic. When two descriptions of the same physical slot carry different classifications, the most restrictive classification is applied to the merged record. Precedence order is: Inclusionary (highest), then Subsidized, then Stabilized, then Locally Unrestricted (lowest). This precedence is enforced at the time of unit record creation, not at query time.

Every rent figure shown to a user carries both a source tier label (Actual, Listed, Calculated, or Estimated) and a specific attribution: the name of the registry, dataset, or method that produced the figure. For Actual and Listed figures aggregated across multiple source records, the attribution names all contributing sources.

5. Coverage and Limitations

5.1 Geographic Scope

The current release covers the City of Mountain View, California. Mountain View is the initial release geography and the primary validation case for the pipeline. The pipeline architecture is designed for multi-jurisdiction expansion; each new release will be accompanied by an updated methodology document and a jurisdiction-specific addendum where the regulatory environment differs materially from Mountain View.

5.2 What Open Rent Observes and Does Not Observe

Asking rent is available for market-rate units with active listings on Dwellsy or RentCast, for income-restricted units as calculated regulatory ceilings via the TCAC method, and for stabilized units at their initial CSFRA registry filing. In-place rent is available for units enrolled in an official filing registry: covered stabilized units (via annual CSFRA updates) and BMR units (via the city's BMR list).

For unregulated market-rate units with no active listing, Open Rent has no rent data of any kind. These units are present in the property inventory and are classified correctly, but no rent figure is shown. The platform does not capture: tenant-specific information, individual lease terms, subletting arrangements, concessions, vacancy rates, or rent figures for properties absent from all ingested source databases.

5.3 Known Coverage Gaps

Five gaps in coverage are documented at the time of this release.

1. **CSFRA non-compliance.** property owners of covered units who have not registered with the City of Mountain View are absent from the rent registry. The registry undercounts the population of stabilized units to a degree that cannot be quantified from available data alone. The ACS tenure and structure-type tables (B25003, B25024) are used as a secondary check on plausible coverage, but cannot resolve this gap directly.
2. **SFR rental identification.** Single-family rental units are identified probabilistically from assessor use codes, parcel characteristics, and listing presence. They are not confirmed rentals. The method is described in the SFR Identification Methodology addendum (see Section 6). False-positive and false-negative rates are documented there.
3. **Listing coverage dependency.** Asking rents for market-rate units are present only when a unit is actively listed on Dwellsy or RentCast at the time of ingestion. Long-term tenanted units, units rented through private channels, and units not indexed by either platform are not represented in the asking rent figures.

4. **Calculated rents are ceilings, not charged rents.** TCAC-derived figures represent the legal maximum allowable rent under the regulatory agreement. Actual charged rents may be lower, particularly where additional local subsidies are layered on top of the tax credit. The discount factors described in Section 4.4 reduce but do not eliminate this gap.
5. **Emerging program sources not yet populated.** HUD Picture of Subsidized Households, CTCAC, and are staged in the pipeline but not yet populated in this release. Their absence means that some federally subsidized units may be present in the property inventory but not fully classified.

5.4 Appropriate and Inappropriate Use

Open Rent data is appropriate for: analyzing the distribution of rental housing stock by regulatory classification; comparing asking rent distributions across building types or ZIP codes; identifying geographic concentrations of stabilized or affordable housing; and tracking changes in the composition of the rental housing stock over time as the dataset is updated.

Open Rent data is not appropriate for: concluding what any individual tenant is currently paying; inferring vacancy rates from the absence of listing data; treating TCAC-calculated rent figures as confirmed charged rents; or drawing conclusions about housing affordability for specific income groups without combining the rent data with independent income distribution data.

6. Additional Methodologies

As Open Rent introduces analytical approaches that go beyond the core data sourcing and processing described in this document (identification models, statistical imputation methods, geographic scoring approaches) those methods are documented as standalone addenda and referenced here. The core methodology remains stable across releases; analytical extensions are versioned independently and published when the feature they support is released.

Each addendum specifies its inputs, method, output format, known limitations, and the version of the platform with which it was released. Where an addendum produces figures that appear in the platform, the platform display includes a link to the relevant addendum.

6.2 Current Addenda

Single-Family Rental Identification Methodology. Describes the scoring model used to identify probable single-family rental units from assessor use codes, parcel characteristics, and listing presence in Mountain View. This model produces the Estimated-tier rent figures described in Section 4.2 for SFR units. It documents the scoring rules, calibration approach, false-positive and false-negative rates, and intended use cases.